Journal Home: https://journals.gmu.edu/index.php/jmms
Mason Journals ● Mason Publishing Group
Journal of Modern Mobility Systems (JMMS), Vol. 03(2022), 105-111
Document: https://doi.org/10.13021/jmms.2022.3494

Comparison of Fixed Targets Utilized by State Agencies and the Use of Varying Targets in the Calibration of Traffic Simulation Models

Md Mahmud Hasan Mamun[a,*], Kamar Ali Amine[a], Mohammed Hadi[a], Syed Ahnaf Morshed[b]

[a] *Florida International University, 10555 West Flagler Street, EC 3730, Miami, FL 33174;* [b] *Institute of Transportation Engineers, 1627 I St. NW, #550, Washington, DC 20006*

Corresponding Author

Received: Dec. 2022 ● **Revised**: 24 Jan. 2023 ● **Accepted**: 31 Jan 2023 ● **Published**: 1 Feb. 2023

Sponsor: https://ATPIO.org

ABSTRACT

Several state departments of transportation have developed simulation guidelines with a strong emphasis on model calibration. These guidelines generally specify fixed targets for each performance measure used in the calibration. However, only a few studies exist that explain the rationale behind and the justification for selecting these calibration targets. In 2019, the Federal Highway Administration (FHWA) released an update to its traffic simulation development methodology, with the concept of calibration targets that vary by segment, time period, and simulated operational scenario depending on the day-to-day variations in the calibrated measure values. As transportation agencies consider the possibility of adopting the FHWA calibration methodology, questions are being raised by the simulation community on how different are the values of the varying targets calculated using the FHWA methodology compared to the existing fixed targets that have been used in the industry for a long time. The purpose of this study is to compare the results from the use of varying calibration targets, as proposed in the methodology of the 2019 guidance, with the fixed targets in the state DOT guidance.

1. **Keywords**: Microsimulation; Traffic Simulation; Traffic Modeling; Traffic Analysis; Calibration Guidelines; Calibration Targets.

1. INTRODUCTION

Calibration is the process of modifying simulation model parameters to enhance a model's capability in emulating time-dynamic system performance observed under particular travel situations [1]. Analysts usually use time-variant macroscopic measures in the calibration process, such as volume, capacity, queue discharge flow, speed, and travel times [2], [3]. The Traffic Analysis Toolbox Volume III, released by the Federal Highway Administration (FHWA) in 2019, provides a detailed traffic simulation development and calibration methodology, which is a revision of the methodology presented in the 2004 version of this document [1], [4]. The 2019 methodology provides a detailed discussion of the use of data for calibration, identifies performance measures and scenarios for modeling based on clustering, selects a representative day for each modeled scenario, and sets calibration targets for each scenario. The methodology in the 2019 version replaced the concept of using fixed calibration targets in the 2004 version with the concept of targets that vary by segment, time period, and simulated operational scenario depending on the day-to-day variations in the calibrated measure values [1], [4]. The methodology provides four criteria for calibration and methods to estimate the targets associated with each criterion. These four criteria are based on the standard deviation of observed data and the deviation between the observed data and simulated data.

Several state departments of transportation (DOTs) have developed simulation guidelines with a strong emphasis on model calibration. These include, for example, the Florida Department of Transportation (FDOT) [5], Iowa Department of Transportation (IDOT) [6], Virginia Department of Transportation (VDOT) [7], [8], Washington State Department of Transportation (WSDOT) [9], Oregon Department of Transportation (ODOT) [10], and Wisconsin Department of Transportation (WisDOT) [11]. Unlike the updated FHWA methodology, the guidance documents of the state DOTs provide fixed thresholds for each MOE used in calibrating the simulation models, and many of them are based on the 2004 FHWA guidance mentioned above.

Only a few studies exist that explain the rationale behind and the justification for selecting calibration targets for the measures in the state guidelines. As transportation agencies consider the possibility of adopting the 2019 FHWA calibration methodology, questions are being raised by the simulation community on how different are the values of the varying targets calculated using the FHWA methodology compared to the existing fixed targets that have been used in the industry for a long time. The purpose of this study is to compare the results from the use of varying calibration targets, as proposed in the methodology in the 2019 guidance, with the fixed targets in the state DOT guidance.

2. LITERATURE REVIEW

Studies found that the absence of proper guidelines for model calibration will lead to incorrect model results [3], [12]. Milam and Choa [12], in an early study, recommended a set of guidelines to calibrate and validate traffic simulation models. Chu et al. [13] provided a multistage and systematic calibration process by considering microscopic driving behavior and route choice model parameters. Dowling et al. [2] suggested a three-step framework for calibrating microsimulation models consisting of calibrating bottleneck capacity, route choice, and system performance. The authors highlighted the calibration targets utilized by WisDOT in a 2002 project, which was based on the guidelines of the Department for Transport, London, United Kingdom [11], [14]. It appears that the United Kingdom guidance is the source of information used in the 2004 version of FHWA guidelines and later in many state agency guidance.

3. STUDY TASKS

A flowchart of the research approach in this study is shown in Figure 1. This study first compared the calibration target setting as recommended in state guidelines. Next, the study examined the method to calculate varying targets in the 2019 update to the FHWA guidelines [4]. The study then compared the use of fixed and varying targets using a case study.

4. CALIBRATION TARGETS IN STATE DOT GUIDELINES

Table 1 summarizes the calibration targets used by state DOTs and indicates that volume, travel time (TT), and speed are the most used measures. Table 1 shows that different states use several types of goodness of fit criteria. Table 1 also shows that the FDOT and IDOT have the same criteria that specify the simulated link volumes for more than 85% of links to be within 100 vehicles per hour (vph) of field measurements for volumes less than 700 vph, within 15% of field measure for volumes between 700 vph and 2700 vph, and within 400 vph of field measure for volumes greater than 2700 vph. The ODOT has the same criterion as the FDOT and IDOT for link volumes higher than 2700 vph but does not specify the criteria for other link volumes. The VDOT uses different targets for volume calibration, and the WisDOT utilizes additional goodness of fit criteria instead of the percent and/or value difference used by other states.

For TT, the FDOT, IDOT, and ODOT have the same thresholds for calibration, and the rest of the states have different thresholds. These three state DOTs specify that the simulated travel times for more than 85% of the network links should be within one minute of the field measurements for links with TT of less than or equal to 7 minutes (min) and ± 15% for links with TT more than 7 min. It should be noted that the FDOT, IDOT, and ODOT utilize values similar to the values presented as example targets in the 2004 version of the FHWA guidelines [4], [11]. The VDOT requires the simulated values to be within 30% of the observed values for arterial streets and within 20% for freeways. The WSDOT uses equations to set the calibration target for TT as a function of real-world TT, segment length, and free-flow speed.

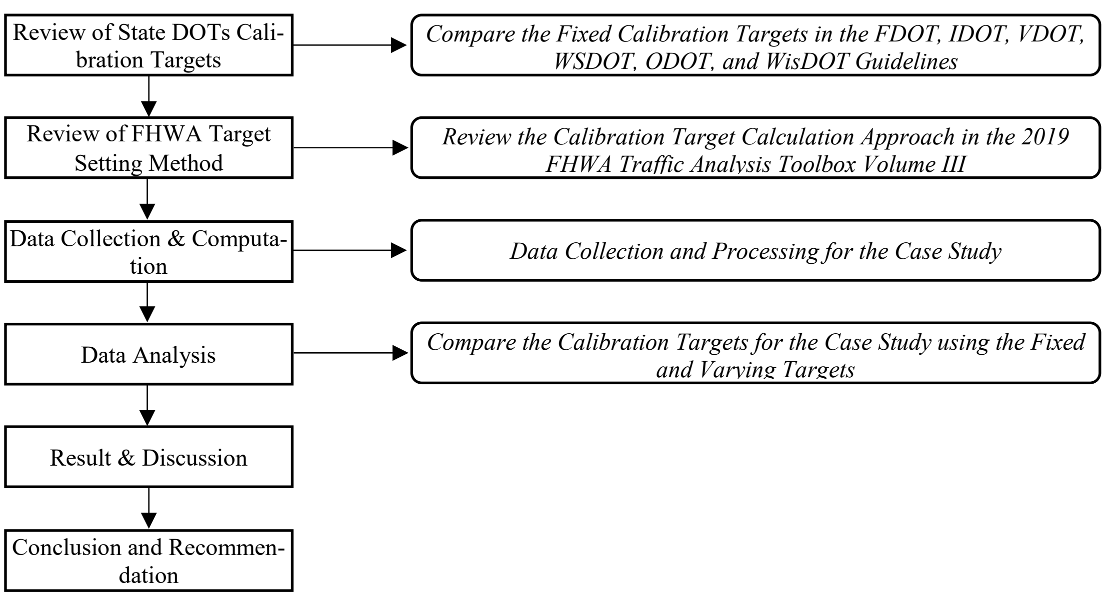

Figure 1. Research Approach

Table 1. Comparison of the Calibration Targets in the Guidance of State Agencies

MOEs	Calibration Targets	FDOT	IDOT	VDOT	WSDOT	ODOT	WisDOT
Traffic Volume	Simulation link Volume to measured field volume for more than 85% of the links of the model	± 100 vph for < 700 vph	± 100 vph for < 700 vph	*±20% for <100 vph*		± 400 vph for > 2700 vph	RMSPE < 5% for > 100 vph (Tier 1)
		± 15% for 700-2700 vph	± 15% for 700-2700 vph	*±15% for 100-1000 vph*			RNSE < 3% for > 100 vph (Tier 2)
		± 400 vph for > 2700 vph	± 400 vph for > 2700 vph	*±10% for 1000-5000 vph*			RNSE < 3% for > 75% of all turns
				± 500 vph for ≥5000 vph			
	GEH for 85% or more simulated links' volume	5 or less	N/A	N/A	< 5 (local roadway)	<5 (Freeway only)	
	Sum of link volumes within calibration area	± 5%	N/A	N/A	± 5%	± 5%	
	GEH for the sum of all link volumes within the calibration area	5 or lower	N/A	N/A	< 3 (all state roadway)		
	GEH at all entry & exit locations and entry & exit ramps within the calibration area	N/A	N/A	N/A	< 3	<5	
Travel Time, TT (includes Transit)	Simulated TT versus observed TT	See the cells below	See the cells below	within 30% of the observed values for arterial streets and 20% for freeways		See the cells below	RMSPE shall be < 10%, and simulated TT shall be within ± 15% for more than 85% of routes with a length greater than 1.5 miles.
	Simulated TT to observed for more than 85% of the network links; Routes with ≤ 7 min TT	± 1 min	± 1 min			± 1 min	
	Simulated TT to observed for more than 85% of the network links Routes with > 7 min TT	±15%	± 15%			± 15%	
Spot Speed (mph)	Average link speed model to field measured for at least 85% of links	±10 mph	±10 mph			± 10 mph	±20%
	Spot speed in the model to field measurement				Uninterrupted flow: ± 3 mph Interrupted flow: ± 10%	± 10%	RMSPE < 10%
Queue Length	Simulation queue length to field	± 20%	±20%	Visually accepted	Qualitative analysis required	Qualitative analysis required	± 150 ft for 300-750 ft queue, ±20% for > 700 ft queue
Lane Use	Lane utilization, model vs. field	Visual Check			Should use lane-by-lane volume and follow volume criteria	GEH ≤ 5	RNSE 3% for >85% of critical lanes

Mason Publishing

5. VARYING CALIBRATION TARGET METHOD

The calibration method in the 2019 FHWA guidelines FHWA [1] involves selecting a representative day for each modeled operational scenario using clustering analysis based on traffic operation parameters such as volumes and travel times, incidents, weather events, and work zones. The guidance then specifies calibration targets as follows:

- About 95% of simulated outputs must fall within a statistical range or two-sigma band range around the representative day values
- Two-thirds of simulated results must fall within a one-sigma band around the representative day values
- The average absolute deviation of simulated measure values from the real-world values of the representative day should be within the target Bounded Dynamic Absolute Error (BDAE) threshold, which reflects the deviation between the representative day and all non-representative days in a group of days that are grouped as one scenario to be modeled in the simulation effort [1].
- The absolute value of the average deviations of each simulated value from the real-world values of the representative day should be less than one-third of the BDAE threshold.

6. DATA COLLECTION AND PROCESSING

The utilized case study investigating the calibration targets is an I-95 facility segment in Broward County and Palm Beach County, Florida, which is 34.3 miles long with 23 interchanges, as shown in Figure 2.

The traffic volume and speed data were collected from an FDOT data archive of traffic sensor measurement. This study segmented the I-95 corridor into subsegments or links that connect the interchange locations.

This study utilized the event-free days after removing weekends and weekdays with precipitation and incidents from the entire year's data. The analysis is first presented using only one scenario for each of the two peak periods, then after clustering the days in multi-scenarios. The representative day for each scenario selected for analysis was obtained based on the representativeness-distance that is calculated as the average of all the distances between each individual day and the most representative day using the equation presented in the FHWA methodology [1].

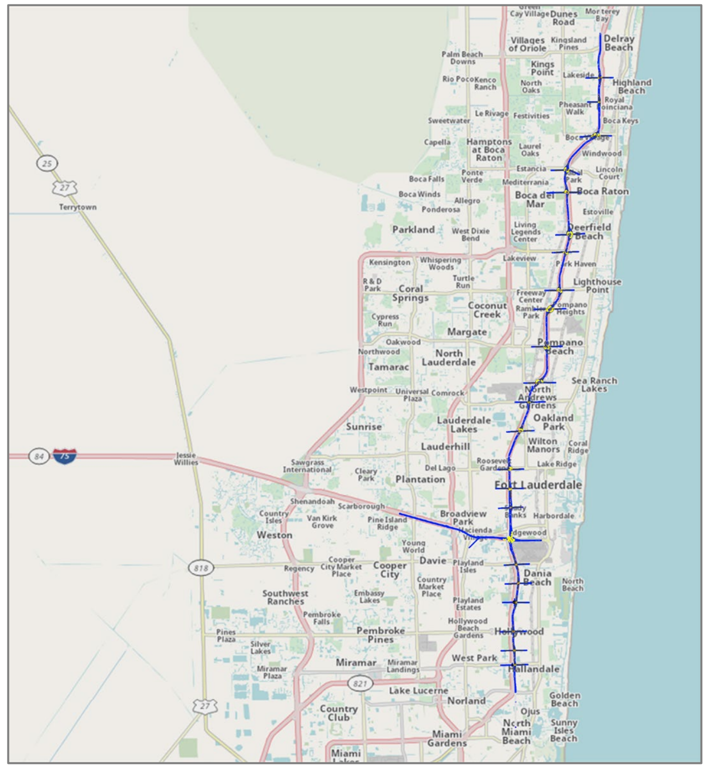

Figure 2. The I-95 Freeway Corridor in Broward and Palm Beach County, Florida

7. RESULTS OF GROUPING ALL EVENT-FREE DAYS IN ONE GROUP

This section presents the results of comparing the calibration targets based on the FHWA methodology (Criteria I and II) with the fixed targets used by FDOT, IDOT, and ODOT.

7.1. Volume Target Comparison

Table 2 shows examples of the results of the volume target comparison for two 15-minute periods. Table 2 indicates that the states' calibration target is, to some extent, comparable to the targets calculated based on the FHWA method for standard deviations lower than 85 vehicles per 15 minutes. When the standard deviation is 100 vehicles per 15 minutes, the state-fixed target is more conservative for links that have standard deviations around this value. With further increases in the standard deviation, the increase of the calculated target above the fixed target increases and reaches unreasonably high values for some links.

Table 2. Comparison of the Results from Applying the Varying Threshold and Fixed Threshold Methods to Calculate Volume Calibration Targets

Volume Measurements (5:00-5:15 pm))						Volume Measurements (6:00-6:15 pm)					
Links	Field Data*, veh/15-min	SD, veh/15-min	2σ (vph, % of volume)	1σ (vph, % of volume)	400 vph (vph, % of volume)	Links	Field Data*, veh/15-min	SD, veh/15-min	2σ (vph, % of volume)	1σ (vph, % of volume)	400 vph (vph, % of volume)
Atlantic Blvd	1074	45	351, 8%	179, 4%	400, 9%	Atlantic Blvd	1322	63	492, 9%	251, 5%	400, 8%
Commercial Blvd	1751	56	441, 6%	225, 3%	400, 6%	SW 10th St	1567	69	538, 9%	274, 4%	400, 6%
Cypress Creek Rd	1817	65	510, 7%	260, 4%	400, 6%	Davie Blvd	1280	76	594, 12%	303, 6%	400, 8%
Davie Blvd	1336	69	543, 10%	277, 5%	400, 7%	I-595	1490	76	597, 10%	305, 5%	400, 7%
I-595	1462	75	590, 10%	301, 5%	400, 7%	Glades Rd	1075	85	664, 15%	339, 8%	400, 9%
Congress Avenue	1735	94	737, 11%	376, 5%	400, 6%	Hillsborough Blvd	1934	106	827, 11%	422, 5%	400, 5%
Copans Rd	1635	98	768, 12%	392, 6%	400, 6%	Commercial Blvd	1871	109	851, 11%	434, 6%	400, 5%
SW 10th St	1516	109	854, 14%	436, 7%	400, 7%	Cypress Creek Rd	1942	112	880, 11%	449, 6%	400, 5%
Linton Blvd	2090	117	917, 11%	468, 6%	400, 5%	Copans Rd	1732	114	893, 13%	456, 7%	400, 6%
Sample Rd	1362	119	931, 17%	475, 9%	400, 7%	Congress Avenue	1669	117	915, 14%	467, 7%	400, 6%
Oakland Park Blvd	1788	122	954, 13%	487, 7%	400, 6%	Sunrise Blvd	1829	130	1016, 14%	519, 7%	400, 5%
Yamato Rd	1462	130	1021, 17%	521, 9%	400, 7%	Yamato Rd	1389	130	1018, 18%	519, 9%	400, 7%
Palmetto Park Rd	981	139	1090, 28%	556, 14%	400, 10%	Palmetto Park Rd	911	134	1048, 29%	535, 15%	400, 11%
Sunrise Blvd	2134	148	1163, 14%	594, 7%	400, 5%	Sample Rd	1195	144	1127, 24%	575, 12%	400, 8%
Glades Rd	969	151	1182, 30%	603, 16%	400, 10%	Oakland Park Blvd	1921	152	1190, 15%	607, 8%	400, 5%
Hillsborough Blvd	1921	216	1692, 22%	863, 11%	400, 5%	Linton Blvd	1926	153	1202, 16%	613, 8%	400, 5%
Broward Blvd	1369	297	2325, 42%	1186, 22%	400, 7%	Broward Blvd	1381	275	2154, 39%	1099, 20%	400, 7%

*Field data reference volume data measured in the identified representative day; SD: Standard Deviation

7.2 Travel Time Targets

Table 3 shows examples of the comparison between the travel time targets calculated according to the FHWA method and the one-minute criterion of the three state agencies. The results in Table 3 show that a total of 76% of the links (13 out of 17) have one sigma less than 35 seconds and two sigma less than 70 seconds. The results indicate that the specification of the one-minute criterion appears to be less conservative when compared to thresholds calculated using the FHWA method, as only the links with higher standard deviation have two-sigma values of more than 60 seconds. A reason for this might be that the examined links were short with short travel times.

Table 3. Comparison of the Results from Applying the Varying Threshold and Fixed Threshold Methods to Calculate Travel Time Calibration Targets

Travel Time Measurements (5:00-5:15 pm)						Travel Time Measurements (6:00-6:15 pm)					
Links	Field Data*, sec	SD, sec	2σ (sec, %)	1σ (sec, %)	1 minute (sec, %)	Links	Field Data*, sec	SD, sec	2σ (sec, %)	1σ (sec, %)	1 minute (sec, %)
Congress Avenue	65	5	10, 16%	5, 8%	60, 93%	Congress Avenue	61	8	16, 27%	8, 14%	60, 98%
Linton Blvd	71	6	11, 16%	6, 8%	60, 84%	Davie Blvd	41	12	24, 59%	12, 30%	60, 145%
Davie Blvd	42	16	31, 73%	16, 37%	60, 143%	SW 10th St	56	16	32, 57%	16, 29%	60, 108%
Yamato Rd	107	17	33, 31%	17, 16%	60, 56%	Linton Blvd	69	18	35, 51%	18, 26%	60, 87%
Cypress Creek Rd	103	17	33, 32%	17, 16%	60, 58%	I-595	86	18	35, 41%	18, 21%	60, 69%
Commercial Blvd	139	19	37, 27%	19, 14%	60, 43%	Copans Rd	58	22	43, 74%	22, 38%	60, 104%
Copans Rd	60	19	37, 62%	19, 32%	60, 99%	Yamato Rd	133	24	46, 35%	24, 18%	60, 45%
I-595	87	23	46, 52%	23, 27%	60, 69%	Hillsborough Blvd	126	27	53, 42%	27, 21%	60, 48%
Atlantic Blvd	105	27	53, 50%	27, 26%	60, 57%	Sample Rd	126	29	57, 45%	29, 23%	60, 48%
SW 10th St	56	27	53, 94%	27, 48%	60, 107%	Palmetto Park Rd	68	30	59, 86%	30, 44%	60, 88%
Oakland Park Blvd	144	28	54, 38%	28, 19%	60, 42%	Commercial Blvd	108	31	61, 56%	31, 29%	60, 56%
Sunrise Blvd	145	32	63, 43%	32, 22%	60, 41%	Cypress Creek Rd	100	32	64, 64%	32, 33%	60, 60%
Sample Rd	155	33	64, 41%	33, 21%	60, 39%	Atlantic Blvd	106	33	66, 62%	33, 31%	60, 56%
Broward Blvd	75	36	70, 94%	36, 48%	60, 80%	Sunrise Blvd	137	35	69, 51%	35, 26%	60, 44%
Glades Rd	245	39	76, 31%	39, 16%	60, 24%	Broward Blvd	71	40	78, 110%	40, 56%	60, 85%
Hillsborough Blvd	130	48	93, 72%	48, 37%	60, 46%	Oakland Park Blvd	96	40	79, 82%	40, 42%	60, 62%
Palmetto Park Rd	114	51	100, 88%	51, 45%	60, 53%	Glades Rd	223	59	115, 52%	59, 26%	60, 27%

*Field data reference volume data measured in the identified representative day; SD: Standard Deviation; sec: seconds.

8. CONCLUSIONS

The study compared a fixed target for volume calibration used by three state agencies with the varying targets calculated using the FHWA methodology for the freeway facility used as a case study. The results indicate that the examined state calibration target is comparable to the calculated targets calculated based on the FHWA method when the standard deviation of the link volume is relatively low (below 85 vehicles per 15 minutes). However, the states' calibration target becomes more conservative when the standard deviation is higher, and this difference increases with the increase in the standard deviation. The study also found that the fixed travel time target appears to be less restrictive compared to the targets calculated using the FHWA methodology for freeway links with relatively short travel times and lower standard deviations, as is in the case study. However, the conclusion could be different for segments with higher travel times and lower standard deviations.

The analysis in this study indicates that assuming that the event-free days belong to one pattern can result in high standard deviations of the MOEs of some of the links, resulting in large acceptable deviations between the simulation results and real-world measurements. Thus, further clustering of the event-free days into more than one pattern is recommended when applying the FHWA methodology, when the standard deviations of the performance measures are high. The impact of this clustering was verified in this study. However, the results are not presented due to the limitation on the size of the paper.

The analysis in this study was conducted using a freeway segment as a case study. Additional research is needed on arterial segments to better understand the applicability of FHWA's approach on arterials. In addition, research is needed to assess the impact of data quality from different sources on the acceptability of the calculated thresholds. Further analysis is needed on how the varying threshold methods impact and are impacted by advanced modeling approaches such as multi-resolution analysis and multi-scenario analysis, including varying demand levels, incidents, and weather conditions.

9. ACKNOWLEDGMENTS

The presented work in this paper is part of a research project funded by the Florida Department of Transportation (FDOT) [Grant number BDV29 977-61]. The opinions, findings, and conclusions expressed in this paper are those of the author(s) and not necessarily those of the funding agencies.

REFERENCES

[1] K. Wunderlich, M. Vasudevan, and P. Wang, "Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software 2019 Update to the 2004 Version," Federal Highway Administration, Washington, DC, 2019. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/43570.

[2] R. Dowling, A. Skabardonis, J. Halkias, G. McHale, and G. Zammit, "Guidelines for Calibration of Microsimulation Models: Framework and Applications," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1876, no. 1, pp. 1–9, Jan. 2004, doi: 10.3141/1876-01.

[3] J. Hourdakis, P. G. Michalopoulos, and J. Kottommannil, "Practical Procedure for Calibrating Microscopic Traffic Simulation Models," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1852, no. 1, pp. 130–139, Jan. 2003, doi: 10.3141/1852-17.

[4] R. Dowling, A. Skabardonis, and V. Alexiadis, "Traffic Analysis Toolbox Volume III : Guidelines for Applying Traffic Microsimulation Modeling Software," U.S. Department of Transportation, Washington, D.C., 2004.

[5] FDOT, *Traffic Analysis Handbook: A Reference for Planning and Operations*, 1st ed., no. March. Tallahassee, FL, USA: Systems Planning Office, Florida Department of Transportation, 2014.

[6] Iowa DOT, *Iowa DOT Microsimulation Guidance*. Iowa City, Iowa: Iowa Department of Transportation, 2017.

[7] VDOT, "VDOT Vissim User Guide," Virginia, USA, 2020.

[8] VDOT Traffic Engineering Division, *Traffic Operations and Safety Analysis Manual*, vol. Version 2. Richmond, VA, USA: Virginia Department of Transportation, 2020.

[9] L. Schilperoort *et al.*, "Protocol for Vissim Simulation WSDOT," Washington, USA, 2014.

[10] C. Mai *et al.*, "Protocol For VISSIM Simulation ODOT," Oregon, USA, 2011.

[11] WisDOT, "Microscopic Simulation Traffic Analysis," in *Traffic Engineering, Operations & Safety Manual*, Wisconsin Department of Transportation, 2019, p. 47.

[12] R. T. Milam and F. Choa, "Recommended Guidelines for the Calibration and Validation of Traffic Simulation Models," 2001, [Online]. Available: https://transportation.ky.gov/Planning/Documents/CORSIM Validation.pdf.

[13] L. Chu, H. X. Liu, J. S. Oh, and W. Recker, "A calibration procedure for microscopic traffic simulation," in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2003, vol. 2, pp. 1574–1579, doi: 10.1109/ITSC.2003.1252749.

[14] Department for Transport London, "Traffic Appraisal of Roads Schemes," in *Design Manual for Roads and Bridges*, London, UK: Department for Transport, 1996, p. 149.